

ne-neu



1600

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003
 TIME: 13:27:00

Input Set : N:\AMC\US09470667.raw
 Output Set: N:\CRF4\02262003\I470667.raw

1 <110> APPLICANT: Asakura, Akira
 2 Hoshino, Tatsuo
 3 Ojima, Setsuko
 4 Shinjoh, Masako
 5 Tomiyama, Noribumi
 6 <120> TITLE OF INVENTION: Novel Alcohol/Aldehyde Dehydrogenases
 7 <130> FILE REFERENCE: C38435/109700CON
 8 <140> CURRENT APPLICATION NUMBER: US/09/470,667
 9 <141> CURRENT FILING DATE: 1999-12-22
 10 <150> PRIOR APPLICATION NUMBER: US 08/934,506
 11 <151> PRIOR FILING DATE: 1997-09-19
 12 <160> NUMBER OF SEQ ID NOS: 12
 13 <170> SOFTWARE: PatentIn version 3.1
 15 <210> SEQ ID NO: 1
 16 <211> LENGTH: 1740
 17 <212> TYPE: DNA
 18 <213> ORGANISM: Gluconobacter oxydans
 19 <400> SEQUENCE: 1
 20 atgaaaccga ctgcgtgct ttggccagt gctggcgac ttgcattgct tgccgcaccc 60
 21 gccttgctc aagtgacccc cgtcaccgat gaattgtgg cgaacccgcg cgctggtaa 120
 22 tggatcagct acggtcagaa ccaagaaaaac taccgtcaact cggccctgac gcagatcag 180
 23 actgagaacg tcggccaact gcaactggtc tggcgccg gcatgcagcc gggcaagtc 240
 24 caagtacgc ccctgatcca tgacggcgatc atgtatctgg caaacccggg cgacgtgatc 300
 25 caggccatcg acgccaaaac tggcgatctg atctggAAC accggcgcca actgccaac 360
 26 atgcgcacgc tgaacagctt tggcgagccg acccgccggca tggcgctgtat cggcaccac 420
 27 gtttactttt tttcggtggca aaccacatcg tgcgcctcg acaccgcaac tggccaagt 480
 28 acgttcgacg tcgaccggcg ccaaggcgaa gacatggttt cgaactcgatc gggcccgatc 540
 29 gtggcaaaacg gctgtatcgatc tggcggttcg acctggcaat actcgccgtt cggctgctt 600
 30 gtctcgggccc acgactcgcc caccgggtgaa gagctgtggc gcaactactt catcccgcc 660
 31 gctggcgaag agggtgatga gacttggggc aacgattacg aagccctgtt gatgaccgg 720
 32 gcctggggcc agatcaccta tgaccccgatc accaacctt tccactacgg ctcgaccgatc 780
 33 gtgggtccgg cggtggaaac ccaacgccc accccggggc gacgctgtat cggcaccac 840
 34 acccgttcg ccgtgcgtcc tgacacgggc gagattgtt ggcgtccatc gaccctggcc 900
 35 cgcgacaact gggaccagga atgcacgttc gagatgtatgg tcaccaatgt gatgtccaa 960
 36 ccctcgaccg agatggaaagg tctgcgtcg atcaacccga acggcgcaac tggcgagcg 1020
 37 cgcgtgctga cccgggttcc gtcggaaaacc ggcaccatgt ggcgtccatc cggcggaaacc 1080
 38 ggcgaattcc tggggcccg tgcataccaaatc taccagaaca tgatcgaatc catcgacgaa 1140
 39 aacggcatcg tgaccgtgaa cgaagatgcg atcctgaagg aactggatgt tgaatatgac 1200
 40 gtctggccga cttcttggg cggccggcgcac tggccgtcg cgcactgaa ccccgacagc 1260
 41 ggcattact tcatacccgatc gaacaacgtc tgctatgaca tgatggccgt cgatcaggaa 1320
 42 ttacacctga tggacgtcta taacaccaggc aacgtgacca agctggccgccc cggcaaggat 1380
 43 atgatcggtc gtattgacgc gatcgacatc agcacgggtc gtacgctgt gtcggctgaa 1440
 44 cgtgctgccc cgaactattc gcccgttcc tgcggccggc gccggcttctt gttcaacgg 1500

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003
TIME: 13:27:00

Input Set : N:\AMC\US09470667.raw
Output Set: N:\CRF4\02262003\I470667.raw

45	ggtaacggatc gttacttccg cgccttcagc caagaaaccg gcgagaccct gtggcagacc	1560
46	cgccttgc aa ccgtcgctc gggccaggcc atctttagc aggttgcggg catgcaat	1620
47	gtcggccatcg cagggtggg tgcgtatgg ggctcggcc tgaactcggc actggctggc	1680
48	gagcggatcg actcgaccgc catcgtaac gccgtctacg tcttcggccct gccgcaataa	1740
50	<210> SEQ ID NO: 2	
51	<211> LENGTH: 1740	
52	<212> TYPE: DNA	
53	<213> ORGANISM: Gluconobacter oxydans	
54	<400> SEQUENCE: 2	
55	atgaagacgt cgtctttgt ggttgcgagc gttgccgc ttgcaagcta tagctcctt	60
56	gcccggctc aagtgcaccc cgtcaccatc gaattgtgg cgaacccggcc cgctggtaa	120
57	tggatcagct acggtcagaa ccaagaaaac taccgtact cgcggctgac gcagatcag	180
58	actgagaacg tcggccact gcaactggc tggcgcgcg gcatgcagcc gggcaaaatgc	240
59	caagtacgc ccctgatcca tgacggcgtc atgtatctgg caaacccggg cgacgtgatc	300
60	caggccatcg acggccaaac tggcgatctg atctggaaac accggccgca actgcccac	360
61	atgcgcacgc tgaacagctt tggcgagccg acccgccgca tggcgctgta cggcaccac	420
62	gttactttg ttcgtggga caaccacctg gtcgcctcg acaccgcaac tggcaagtg	480
63	acgttcgacg tcgaccggcg ccaaggcgaa gacatgtt cgaactcgatc gggcccgatc	540
64	gtggcaaaacg gcgtgatcg tgcgggttcg acctgccaat actgcgcgtt cggctgctt	600
65	gtctcgcc acgactcgcc caccggtaa gagctgtggc gcaactactt catccgcgc	660
66	gctggcgaag agggtgatga gacttggggc aacgattacg aagccgttg gatgaccggc	720
67	gtctgggtc agatcaccta tgaccccggtt ggcggccttg tccactacgg ctcgtcggt	780
68	gttggcccg cttcgaaac ccacgcgcg accacccggc gcaccatgta cggcaccac	840
69	acccggttcg ctgtccgtcc cgagactggc gagatcgatc ggcgtcacca aactctgccc	900
70	cgcgacaact gggaccaaga gtgcacccctc gagatgtatgg ttgccaacgt tgacgtgcag	960
71	cccgacgtc acatggacgg cgtccgtcg atcaacccga acggccgac cggcggcgat	1020
72	cgcgtctga cccggcttcc gtgcacccacc ggcaccatgt ggcgttgcg cggccaaacc	1080
73	ggcgaattcc tggggcccg tgacaccatc tacgagaaca tcatgaaatc gatcgacgaa	1140
74	aacggcatcg tgaccgtcg ctagtcgaaa gttctgaccg agctggacac cccctatgac	1200
75	gtctggccgc tgctgtggg tggccgtgac tggccgtcg ctgcgtgaa ccccgatacc	1260
76	ggcatctact ttatcccgat gaaacaacacc tgcgtatcgata tgcgtatcgatc cgaccaggaa	1320
77	ttcagctcg tggacgtgtt caacccaaacg ctgaccgca aatggcacc gggtaaagag	1380
78	ctgggtggcc gtatcgacgc catcgacatc agcacaggcc gcaccctgtg gaccgttag	1440
79	cgcgaaacgc cgaactacgc gcctgtctcg tgcggctg gccggcttct gttcaacggc	1500
80	ggcaccgacc gttacttccg cgtctcagc caagagaccg gcgagaccct gttccacggc	1560
81	cgtctggcgat ctgtcgatcc gggccaaatgt gtcgtatcg agatcgacgg cgtccaaatc	1620
82	atcgccatcg cggccggcg caccgatctt gttcgatcc acaaccgtcc cctggccgag	1680
83	ccggtcact cgaccgcgat cggtatcgatc atgtacgtt tgcgtcgcc ccagcaataa	1740
85	<210> SEQ ID NO: 3	
86	<211> LENGTH: 1737	
87	<212> TYPE: DNA	
88	<213> ORGANISM: Gluconobacter oxydans	
89	<400> SEQUENCE: 3	
90	atgaaactga cgaccgtct gcaacgcgcg gcccgcgtc ttgtgtttgg caccattccc	60
91	gcccggccaa acccgccat caccgtgaa atgctggcgat acccgccgc tggtaatgg	120
92	atcaactacg gtcagaacca agagaactac cgcactcgcc ccctgacgc gattaccgc	180
93	gacaacgtcg cgcactcgatc actggctgg ggcgcggta tggaaacggg caagatccaa	240
94	gtgacccgcg tttgtccatga cggcgtatcg tatctggca accccggta cgtatccatc	300
95	cccatcgacg cggcggccgg cgatctgatc tggaaacacc gccgccaact gccgacatc	360

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003

TIME: 13:27:00

Input Set : N:\AMC\US09470667.raw

Output Set: N:\CRF4\02262003\I470667.raw

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003
TIME: 13:27:00

Input Set : N:\AMC\US09470667.raw
Output Set: N:\CRF4\02262003\I470667.raw

146 accggcattt acttcttgcc gctgaacaat gcctgctacg atattatggc cgttgatcaa 1320
 147 gagtttagcg cgctcgacgt ctataacacc agcgcgaccg caaaactcgc gccgggcttt 1380
 148 gaaaatatgg gccgcatcga cgcgattgat atcagcaccg ggcgcacctt gtggcggcg 1440
 149 gagcgcctg cggcgaacta ctgcgcgtt ttgtcgacgg caggcgggtgt ggtgttcaac 1500
 150 ggcgggaccg accgctattt ccgtgcctc agccagggaaa cccgcgagac tttgtggcag 1560
 151 gcccgtcttgc acacggtcgc gacggggcag gcgatcagct acgagttgaa cggcgtgcaa 1620
 152 tatatcgcca tcggtgcggg cgtctgacc tatggcacgc aattgaacgc gccgctggcc 1680
 153 gaggcaatcg attcgaccc ggtcgtaat gcgatctatg tctttgcact gccgcagtaa 1740
 155 <210> SEQ ID NO: 5
 156 <211> LENGTH: 579
 157 <212> TYPE: PRT
 158 <213> ORGANISM: Gluconobacter oxydans
 159 <220> FEATURE:
 160 <221> NAME/KEY: SIGNAL
 161 <222> LOCATION: (1)..(23)
 162 <223> OTHER INFORMATION:
 W--> 163 <400> 5
 164 Met Lys Pro Thr Ser Leu Leu Trp Ala Ser Ala Gly Ala Leu Ala Leu
 165 1 5 10 15
 166 Leu Ala Ala Pro Ala Phe Ala Gln Val Thr Pro Val Thr Asp Glu Leu
 167 20 25 30
 168 Leu Ala Asn Pro Pro Ala Gly Glu Trp Ile Ser Tyr Gly Gln Asn Gln
 169 35 40 45
 170 Glu Asn Tyr Arg His Ser Pro Leu Thr Gln Ile Thr Thr Glu Asn Val
 171 50 55 60
 172 Gly Gln Leu Gln Leu Val Trp Ala Arg Gly Met Gln Pro Gly Lys Val
 173 65 70 75 80
 174 Gln Val Thr Pro Leu Ile His Asp Gly Val Met Tyr Leu Ala Asn Pro
 175 85 90 95
 176 Gly Asp Val Ile Gln Ala Ile Asp Ala Lys Thr Gly Asp Leu Ile Trp
 177 100 105 110
 178 Glu His Arg Arg Gln Leu Pro Asn Ile Ala Thr Leu Asn Ser Phe Gly
 179 115 120 125
 180 Glu Pro Thr Arg Gly Met Ala Leu Tyr Gly Thr Asn Val Tyr Phe Val
 181 130 135 140
 182 Ser Trp Asp Asn His Leu Val Ala Leu Asp Thr Ala Thr Gly Gln Val
 183 145 150 155 160
 184 Thr Phe Asp Val Asp Arg Gly Gln Gly Glu Asp Met Val Ser Asn Ser
 185 165 170 175
 186 Ser Gly Pro Ile Val Ala Asn Gly Val Ile Val Ala Gly Ser Thr Cys
 187 180 185 190
 188 Gln Tyr Ser Pro Phe Gly Cys Phe Val Ser Gly His Asp Ser Ala Thr
 189 195 200 205
 190 Gly Glu Glu Leu Trp Arg Asn Tyr Phe Ile Pro Arg Ala Gly Glu Glu
 191 210 215 220
 192 Gly Asp Glu Thr Trp Gly Asn Asp Tyr Glu Ala Arg Trp Met Thr Gly
 193 225 230 235 240
 194 Ala Trp Gly Gln Ile Thr Tyr Asp Pro Val Thr Asn Leu Val His Tyr
 195 245 250 255

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003
TIME: 13:27:00

Input Set : N:\AMC\US09470667.raw
Output Set: N:\CRF4\02262003\I470667.raw

196 Gly Ser Thr Ala Val Gly Pro Ala Ser Glu Thr Gln Arg Gly Thr Pro
 197 260 265 270
 198 Gly Gly Thr Leu Tyr Gly Thr Asn Thr Arg Phe Ala Val Arg Pro Asp
 199 275 280 285
 200 Thr Gly Glu Ile Val Trp Arg His Gln Thr Leu Pro Arg Asp Asn Trp
 201 290 295 300
 202 Asp Gln Glu Cys Thr Phe Glu Met Met Val Thr Asn Val Asp Val Gln
 203 305 310 315 320
 204 Pro Ser Thr Glu Met Glu Gly Leu Gln Ser Ile Asn Pro Asn Ala Ala
 205 325 330 335
 206 Thr Gly Glu Arg Arg Val Leu Thr Gly Val Pro Cys Lys Thr Gly Thr
 207 340 345 350
 208 Met Trp Gln Phe Asp Ala Glu Thr Gly Glu Phe Leu Trp Ala Arg Asp
 209 355 360 365
 210 Thr Asn Tyr Gln Asn Met Ile Glu Ser Ile Asp Glu Asn Gly Ile Val
 211 370 375 380
 212 Thr Val Asn Glu Asp Ala Ile Leu Lys Glu Leu Asp Val Glu Tyr Asp
 213 385 390 395 400
 214 Val Cys Pro Thr Phe Leu Gly Gly Arg Asp Trp Pro Ser Ala Ala Leu
 215 405 410 415
 216 Asn Pro Asp Ser Gly Ile Tyr Phe Ile Pro Leu Asn Asn Val Cys Tyr
 217 420 425 430
 218 Asp Met Met Ala Val Asp Gln Glu Phe Thr Ser Met Asp Val Tyr Asn
 219 435 440 445
 220 Thr Ser Asn Val Thr Lys Leu Pro Pro Gly Lys Asp Met Ile Gly Arg
 221 450 455 460
 222 Ile Asp Ala Ile Asp Ile Ser Thr Gly Arg Thr Leu Trp Ser Val Glu
 223 465 470 475 480
 224 Arg Ala Ala Ala Asn Tyr Ser Pro Val Leu Ser Thr Gly Gly Val
 225 485 490 495
 226 Leu Phe Asn Gly Gly Thr Asp Arg Tyr Phe Arg Ala Leu Ser Gln Glu
 227 500 505 510
 228 Thr Gly Glu Thr Leu Trp Gln Thr Arg Leu Ala Thr Val Ala Ser Gly
 229 515 520 525
 230 Gln Ala Ile Ser Tyr Glu Val Asp Gly Met Gln Tyr Val Ala Ile Ala
 231 530 535 540
 232 Gly Gly Gly Val Ser Tyr Gly Ser Gly Leu Asn Ser Ala Leu Ala Gly
 233 545 550 555 560
 234 Glu Arg Val Asp Ser Thr Ala Ile Gly Asn Ala Val Tyr Val Phe Ala
 235 565 570 575
 236 Leu Pro Gln
 238 <210> SEQ ID NO: 6
 239 <211> LENGTH: 579
 240 <212> TYPE: PRT
 241 <213> ORGANISM: Gluconobacter oxydans
 242 <220> FEATURE:
 243 <221> NAME/KEY: SIGNAL
 244 <222> LOCATION: (1)..(23)
 245 <223> OTHER INFORMATION:

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/470,667

DATE: 02/26/2003

TIME: 13:27:01

Input Set : N:\AMC\US09470667.raw

Output Set: N:\CRF4\02262003\I470667.raw

L:163 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:0
L:246 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:6,Line#:0
L:329 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:0
L:412 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:8,Line#:0
L:511 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:11,Line#:0